Learnable Descriptive Convolutional Network for Face Anti-Spoofing

Pei-Kai Huang[1], Hui-Yu Ni[1], Yan-Qin Ni[2] and Chiou-Ting Hsu[1]

[1]National Tsing Hua University, Taiwan ; [2]National Central University, Taiwan

Face Anti-Spoofing (FAS)

Goal

- **Learning discriminative features**
  - ➢ From visually similar live and spoof faces?
- **Achieving good cross-domain testing performance**
  - ➢ How to narrow the domain gap?

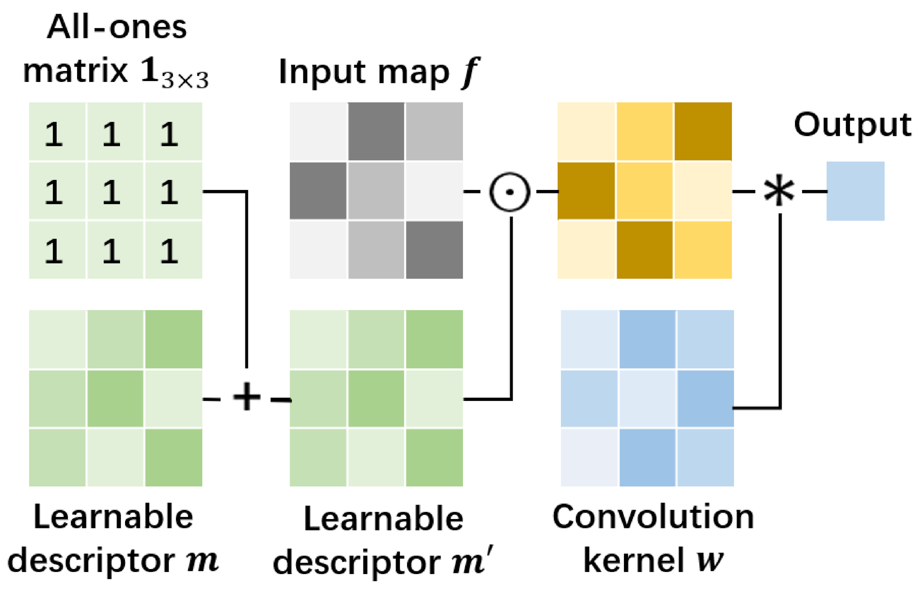
LDCNet: Learnable Descriptive Convolutional Network

Idea

- **Not limited to pre-defined local descriptors**

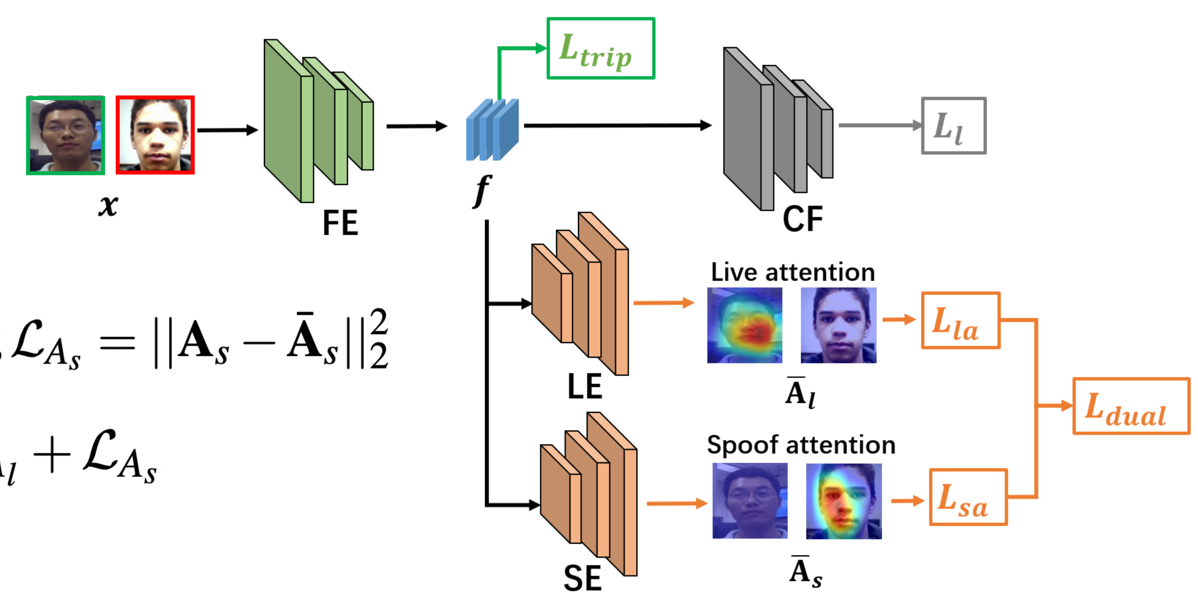
Proposed method

- **Learnable Descriptive Convolution**



$$g(p) = (1-\varepsilon)\underbrace{\sum_{p_n\in\mathcal{R}} w(p_n)\cdot f(p+p_n)}_{\text{vanilla convolution}} + \varepsilon \underbrace{\sum_{p_n\in\mathcal{R}} w(p_n)\cdot(f(p+p_n)\cdot m(p_n))}_{\text{learnable descriptive convolution}}$$

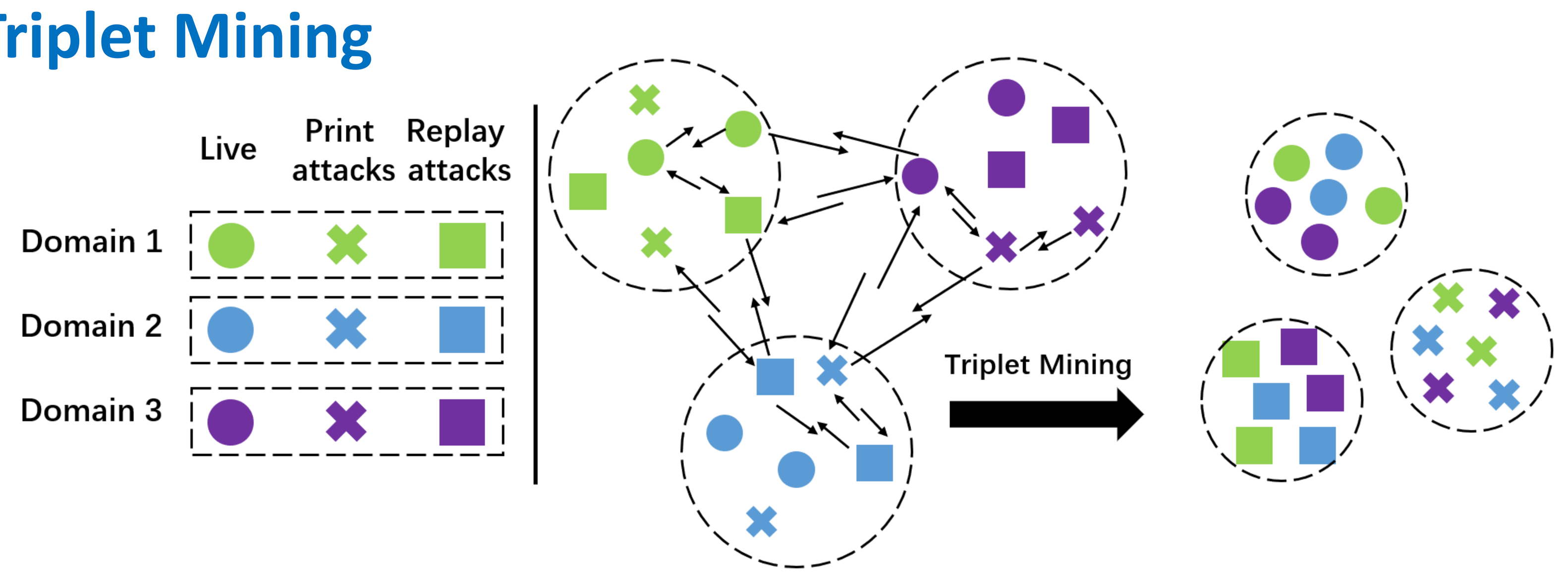
- **Dual Attention Supervision**



$$\mathcal{L}_{A_l} = ||\mathbf{A}_l - \bar{\mathbf{A}}_l||_2^2; \mathcal{L}_{A_s} = ||\mathbf{A}_s - \bar{\mathbf{A}}_s||_2^2$$

$$\mathcal{L}_{dual} = \mathcal{L}_{A_l} + \mathcal{L}_{A_s}$$

- **Triplet Mining**



$$\mathcal{L}_{trip} = \sum_{\forall x_i^a} (||\mathbf{FE}(x_i^a) - \mathbf{FE}(x_i^p)||_2^2 - ||\mathbf{FE}(x_i^a) - \mathbf{FE}(x_i^n)||_2^2 + \alpha)$$

- **Total Loss**

$$\mathcal{L}_T = \mathcal{L}_l + \beta\mathcal{L}_{trip} + \gamma\mathcal{L}_{dual}$$

Datasets

OULU-NPU (O), MSU-MFSD (M), CASIA-MFSD (C), Replay-Attack (I)

Ablation Studies

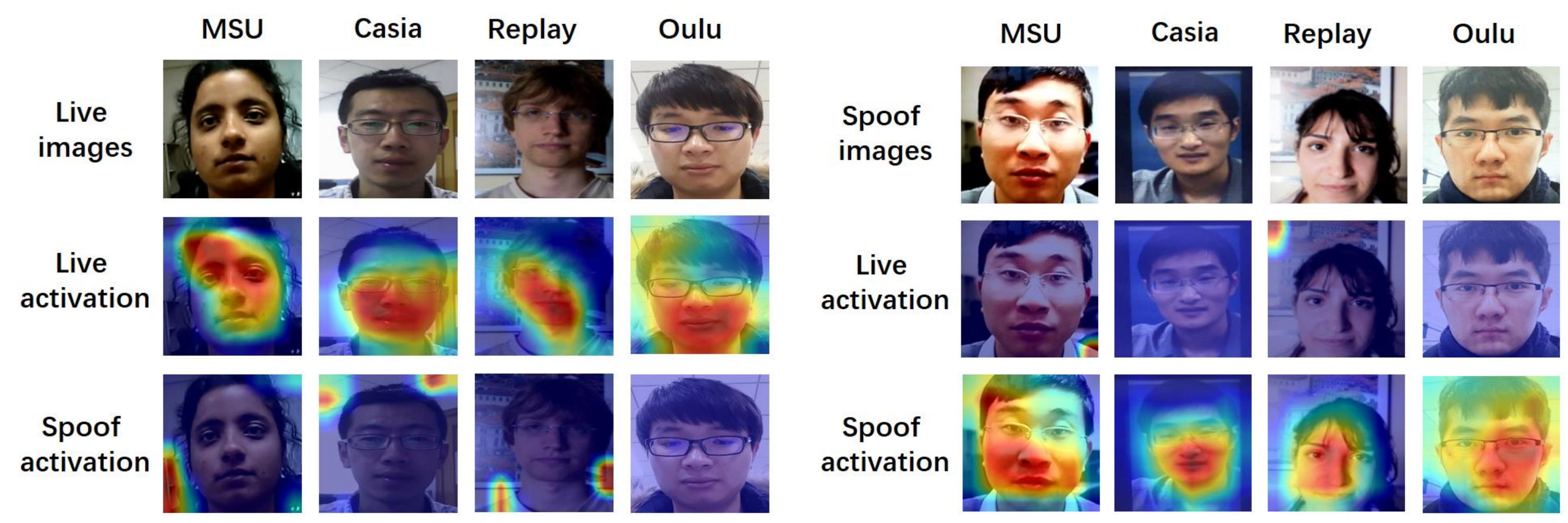
Different combinations of loss terms

Total loss $\mathcal{L}_T$				[O,C,I]→ M		[O,M,I]→ C		[O,C,M]→ I		[I,C,M]→ O	
$\mathcal{L}_l$	$\mathcal{L}_{trip}$	$\mathcal{L}_{dual}$: $\mathcal{L}_{A_l}$	$\mathcal{L}_{dual}$: $\mathcal{L}_{A_s}$	HTER	AUC	HTER	AUC	HTER	AUC	HTER	AUC
✓				15.24	90.43	18.33	89.45	17.14	89.84	18.22	88.37
✓	✓			13.10	91.98	16.67	90.98	13.87	92.84	16.19	90.56
✓	✓	✓		11.43	93.68	13.07	93.89	10.64	94.92	14.32	92.54
✓	✓		✓	12.38	93.14	14.72	92.57	12.71	94.13	14.91	91.39
✓	✓	✓	✓	**9.29**	**96.86**	**12.00**	**95.67**	**9.43**	**95.02**	**13.51**	**93.68**

Different convolution kernels

Method	HTER	AUC
Vanilla Convolution [25]	16.65	84.19
Sobel Convolution [37] (*CVPR 20*)	14.96	90.00
Local Binary Convolution [15] (*CVPR 17*)	15.10	90.50
Central Difference Convolution [42] (*TPAMI 20*)	14.94	91.33
Ours: LDC	**13.51**	**93.68**

Experimental Comparisons



Intra-domain testing

Method	P.	ACPER	BPCER	ACER	P.	ACPER	BPCER	ACER
SGTD [37] (*CVPR 20*)	1	2.0	0.0	1.0	3	3.2±2.0	2.2±1.4	2.7±0.6
BCN [41] (*ECCV 20*)		0.0	1.6	0.8		2.8±2.4	2.3±2.8	2.5±1.1
Disentangle [34] (*CVPR 20*)		1.7	0.8	1.3		2.8±2.2	1.7±2.6	2.2±2.2
RAEDFL [12] (*ACPR 21*)		1.67	0.00	0.83		1.38±1.78	0.28±0.68	**0.83±0.86**
Structure [45] (*IJCB 21*)		1.3	0.0	**0.6**		2.3±2.7	1.4±2.6	**1.9±1.6**
Ours		0.0	0.0	**0.0**		4.55±4.55	0.58±0.91	2.57±2.67
SGTD [37] (*CVPR 20*)	2	2.5	1.3	1.9	4	6.7±7.5	3.3±4.1	5.0±2.2
BCN [41] (*ECCV 20*)		2.6	0.8	1.7		2.9±4.0	7.5±6.9	5.2±3.7
Disentangle [34] (*CVPR 20*)		2.7	2.7	2.4		5.4±2.9	3.3±6.0	4.4±3.0
RAEDFL [12] (*ACPR 21*)		0.69	1.67	**1.18**		5.41±6.40	2.50±2.74	3.96±3.90
Structure [45] (*IJCB 21*)		2.2	2.2	2.2		6.7±6.8	0.0±0.0	**3.3±3.4**
Ours		0.8	1.0	**0.9**		4.50±1.48	3.17±3.49	**3.83±2.12**

Cross-domain testing

Method	[O,C,I]→ M		[O,M,I]→ C		[O,C,M]→ I		[I,C,M]→ O	
	HTER	AUC	HTER	AUC	HTER	AUC	HTER	AUC
MADDG [30] (*CVPR 19*)	17.69	88.06	24.50	84.51	22.19	84.99	27.89	80.02
DR-MD-Net [34] (*CVPR 20*)	17.02	90.10	19.68	87.43	20.87	86.72	25.02	81.47
SSDG-M [13] (*CVPR 20*)	16.67	90.47	23.11	85.45	18.21	94.61	25.17	81.83
RFM [31] (*AAAI 20*)	13.89	93.98	20.27	88.16	17.30	90.48	16.45	91.16
RAEDFL [12] (*ACPR 21*)	16.67	87.93	17.78	86.11	14.64	85.64	18.06	90.04
ANRL [19] (*ACM MM 21*)	10.83	96.75	17.83	89.26	16.03	91.04	15.67	91.90
D^{2}AM [5] (*AAAI 21*)	15.43	91.22	12.70	95.66	20.98	85.58	15.27	90.87
SDA [35] (*AAAI 21*)	15.40	91.80	24.50	84.40	15.60	90.10	23.10	84.30
CDCN-PS [43] (*TBBIS 21*)	20.42	87.43	18.25	86.76	19.55	86.38	15.76	92.43
FAS-DR-BC(MT) [27] (*TPAMI 22*)	11.67	93.09	18.44	89.67	11.93	94.95	16.23	91.18
LMFD-PAD [7] (*WACV 22*)	10.48	94.55	12.50	94.17	18.49	84.72	**12.41**	**94.95**
SSAN-M [38] (*CVPR 22*)	10.42	94.76	16.47	90.81	14.00	94.58	19.51	88.17
SSAN-R [38] (*CVPR 22*)	**6.67**	**98.75**	**10.00**	**96.67**	**8.88**	**96.79**	13.72	93.63
Ours	**9.29**	**96.86**	**12.00**	**95.67**	**9.43**	**95.02**	**13.51**	**93.68**

Limited cross-domain testing

Method	[M,I]→ C		[M,I]→ O	
	HTER	AUC	HTER	AUC
MADDG [30] (*CVPR 19*)	41.02	64.33	39.35	65.10
DR-MD-Net [34] (*CVPR 20*)	31.67	75.23	34.02	72.65
SSDG-M [13] (*CVPR 20*)	31.89	71.29	36.01	66.88
RFM [31] (*AAAI 20*)	36.34	67.52	29.12	72.61
RAEDFL [12] (*ACPR 21*)	31.11	72.63	29.23	74.62
ANRL [19] (*ACM MM 21*)	31.06	72.12	30.73	74.10
D^{2}AM [5] (*AAAI 21*)	32.65	72.04	**27.70**	75.36
SDA [35] (*AAAI 21*)	32.17	72.79	28.90	73.33
SSAN-M [38] (*CVPR 22*)	**30.00**	**76.20**	29.44	**76.62**
Ours	**22.22**	**82.87**	**21.54**	**86.06**

BMVC 2022

The 33rd **British Machine Vision Conference**

21 - 24 November 2022, London, UK